Possibility and prevention of data tampering in the referee process of data journal

Takeshi Terui[1]*, ***Yasuyuki Minamiyama***[1], ***Kazutsuna Yamaji***[2]

[1]* *National Institute of Polar Research, 10-3, Midori-cho, Tachikawa-shi, Tokyo 190-8518, Japan*
[2] *National Institute of Informatics, 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430, Japan*
Email: terui.takeshi@nipr.ac.jp

Summary. The risk of data tampering exists on the peer review process in the data journal. When the Polar Data Journal was launched, we analyzed this possibility and preventive measures. There is a possibility that data tampering is done under the review process whether it is fault or not. What is most important to the data journal is whether the authenticity of the data at posting and publication is preserved. Therefore, Polar Data Journal calculates the hash value of data immediately after posting and immediately before publication, and detects data falsification by comparing hash values. It is expected that this method can easily prevent data tampering.

Keywords. Data Tampering, Review Process, Hash Value, Security.

1. Introduction

In the data journal, the submitted data itself is reviewed. The correctness and validity of the data are evaluated by the editorial board and reviewer than the description contents on the paper. Because data is only a simple digital information asset, there is a risk of information security. The risk that you must pay attention to is the data tampering. Data tampering significantly reduces the trustworthiness of data journals. It is very important to take this preventive measure. This paper introduces measures to prevent the above data falsification in Polar Data Journal.

2. Data Tampering

Data tampering in the data journal is one of the incidents that become apparent when the data published after publication is different from the peer-reviewed data. This incident can happen regardless of purpose or negligence. In the data journal, published data is subject to peer review by experts. However, when data falsification becomes apparent, the peer review process will be overwhelmed. Ultimately, the data journal is considered untrusted and has the property of retreating open science. Therefore, efforts must be made to prevent data falsification.

There is a hash value comparison as a technology for detecting data tampering. In Polar Data Journal, hash values of data at posting and reposting are calculated and stored in the paper submission system. Calculate the hash value of the data released just before publication of the paper again and compare it with the stored hash value. When data alteration is performed, the hash value has different results. The office of Polar Data Journal detects data alteration and reports to editorial committee.

In this research presentation, we introduce a process that is highly likely to be falsified and measures to prevent it.